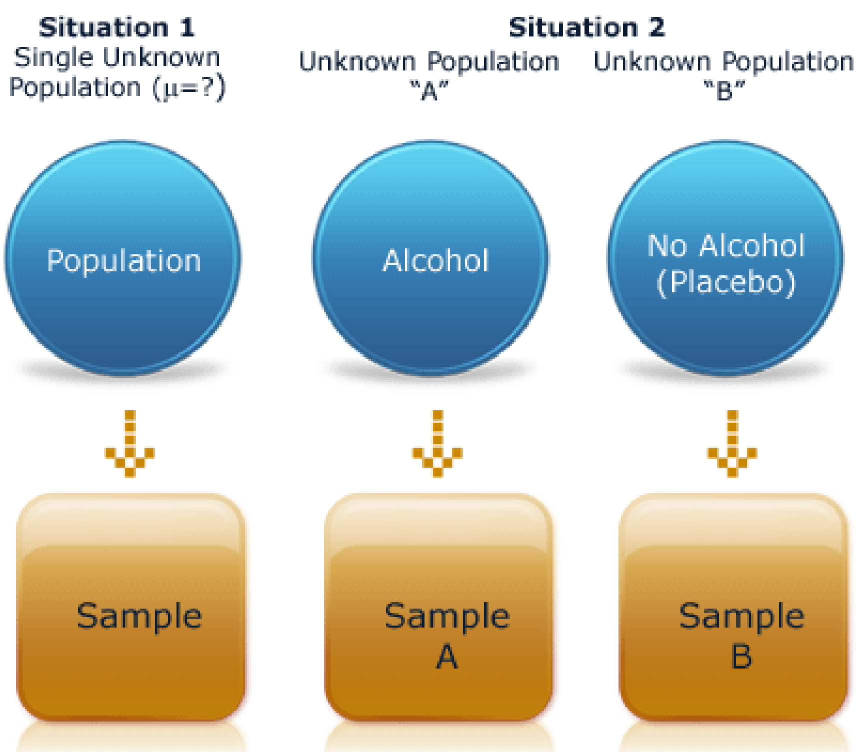
Print

Lesson 14: Dependent Means Hypothesis Testing - Study Notes

Slide 1:

Independent vs. Dependent Samples

The previous sections dealt with the study of samples in order to make inferences about their population. For example, testing the sample mean gave us a hint as to what the population mean should be, with certain confidence. Now our attention has turned to making comparisons between two populations via samples from both.



Situation 1 would represent the type of problems we encountered in the previous lesson, whereas this lesson will focus on how to approach Situation 2. This particular schema seems to have been created for a study involving alcohol consumption.

When making inferences for two populations, we will encounter two different types of samples. Each situation will require its own procedure to carry out the testing.

When the same set of sources is used to obtain the data representing both populations, we have a situation involving **dependent samples**. Using an experimental procedure known as matched-subjects design, dependent samples are paired and compared to each other. When dealing with this type of study, we examine the mean differences between each paired data entry.

Independent samples come from two unrelated sets of sources. This situation considers the difference between the means of the two sample sets in order to make inferences about the populations they came from. Sample sizes do not have to be equal in order to perform this test since they aren't

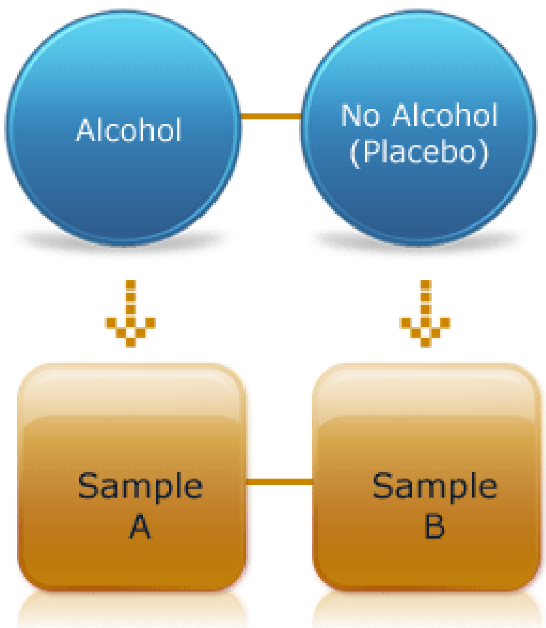
paired.

Slide 2:

Dependent Means Testing

Due to their relationship, dependent samples are paired to each other. The main reason behind this method is to eliminate error that would otherwise be uncontrollable. For example, let's say that we were interested in testing the effects of alcohol consumption on the pain. The hypothesis is that the higher one's blood alcohol level, the less pain they feel (until the next morning!). In order to eliminate confounders such as one's tolerance for pain, a matched-subjects design would be employed such that it is the same individual who is representing two distinct populations. Subjects could first be asked to rate their perception of pain on a 10-point scale after having a needle prick their hand. Then the same subjects would be given the treatment (several alcoholic beverages) and put through the same procedure again.

To do this, individual values from one sample are paired with members of the other so that we can investigate differences in the "paired data.". In other words, we are interested in the difference between the "couple," or the **paired difference**. If we take a closer look at our diagram, we would notice that this is obviously a study involving two different populations. However, are they independent or dependent? Determining the nature of the samples constitutes the first step in this new kind of hypothesis testing.



If the individuals involved in sample A were the same individuals tested in sample B, then we would have a dependent sample. If sample A and B are made up of two different randomly-sampled members, then we're looking at an independent sample.

In our alcohol-pain experiment, the fact that the same subjects are measured twice is a perfect example of a dependent means test. However, had the individuals been split into two separate groups where one is given the treatment (drinking booze) and the other is not (drinking a non-alcoholic substitute), and their tolerance levels were compared, this would be an example of an independent means test.

Let's try an example of a dependent sample and take you through the steps.

Slide 3:

Dependent Means Testing

Example

You are interested in determining the effects of added weight to your car tires. You're a real cheapskate when it comes to spending on your car and you've been telling your mechanics to "leave you alone" when they insist that your vehicle needs a tire rotation (left switched to right, and vice versa). Believing that they are just trying to con you into dishing out some more of your hard-earned money, you decide to use your fat government grant to explore the problem, "for the good of car owners everywhere." (Boy, are you ever bitter!)

- The null hypothesis should be $\mu_d = 0$ (meaning that there is no **difference** between the two sides)

- The alternate hypothesis: $\mu_d \neq 0$

You randomly select 6 vehicles, equip them with brand new tires, and ask the drivers to navigate solo for 500 km a week for 4 weeks. At the end of this period, you measure the amount of wear on each front tire (one left and one right).

Here are the results:

Tire Groove Thickness (thousandths of a cm)

Tire/Car	1	2	3	4	5	6
Left	125	64	94	38	90	106
Right	133	65	103	37	102	115

Now we can compare the individual means by finding the difference between each set of paired data. This is the tricky part. When doing these calculations, we have to keep in mind what we are trying to examine. Our Ho is suggesting that there is no mean difference, meaning that Ha would deal with **any** difference between the tires: positive or negative. This is a two-sided test. In this case, it doesn't matter if we subtract the right tire from the left or vice versa, as long as we stick to the same format all the way through.

Ho: $\mu_d = 0$
Ha: $\mu_d \neq 0$

where "d" stands for the difference between paired values

Slide 4:

Dependent Means Testing (Cont'd)

Example

But let's say, on the other hand, that we wanted to prove that the left tire wore out **more** than the right one. That means that if we subtracted the left tire thickness from the right one, or vice

versa, our values would tend to be either all negative or all positive. More wear on the left tire would mean that R-L would give us a positive number because there is more rubber remaining on the right tire than on the left one (R > L).

If this was the case, we must also alter our hypothesis to match our new theory:

Ho: $\mu_d \leq 0$ Ha: $\mu_d > 0$	OR	Ho: $\mu_d \geq 0$ Ha: $\mu_d < 0$
where "d" is R - L		where "d" is L - R

Again, we want to be very careful as to how we calculate the differences between the tires. If our claim were true (that the left tires wear out more), then subtracting the left tire thickness from the right one would yield a positive number (R - L > 0), which is exactly what our Ha is implying in the first scenario. If we were to go with it, then our calculations would look like this:

Tear wear differences using R - L

Car	**1**	**2**	**3**	**4**	**5**	**6**
X_d	8	1	9	-1	12	9

The average difference between the two tires on each car will be denoted as "d". Therefore, the average difference between all the differences can be calculated as follows:

$$\bar{d} = \frac{\sum d}{n}$$

Using n - 1 as our degrees of freedom (where n stands for the number of paired data), we can calculate the standard deviation of "d" (s_d) using this formula:

$$S_d = \sqrt{\frac{\sum d^2 - \left[\frac{(\sum d)^2}{n}\right]}{n-1}}$$

Note: Notice that the formula for $\bar{d}$ and s_d are exactly the same as for $\bar{x}$ and s, but "x" has been replaced by "d" to signify that we are using the differences in each pair and not the raw scores.

That said and done, let's apply it to our values:

- Average difference between right and left tire thickness (R - L) = $\bar{d}$ = (8+1+9-1+12+9) / 6 = 38 / 6 = **6.33**
- Substituting $\sum d = 38$ and $\sum d^2 = (372)$ into the standard deviation equation for S_d, we get: sqrt (372-240.667)/5 = sqrt (26.267) = **5.13.**

Slide 5:

Dependent Means Testing (Cont'd)

Example

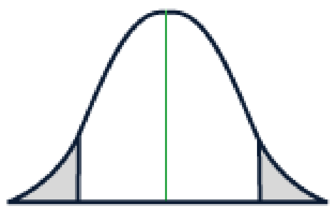
As far as the hypothesis testing goes, recall from the previous lesson that if the sample size was small (less than 30) and the population standard deviation was unknown, we would have to use the t-statistic. Nothing has changed. It is VERY rare that you will encounter a problem of this type where the sample space is large. The reason for this is that it is very time-consuming to find paired differences for a large amount of data values...and it takes even longer to correct!

In our example, since our sample size is small (it's 6 since we've investigated 6 cars or 6 pairs of tires, not 12 tires!), and we are not given the population standard deviation, we must use the t-stat. For inferences involving dependent means, the formula looks like this:

$$t^* = \frac{\bar{d}-\mu_d}{s_d/\sqrt{n}}$$

As you can see, this formula is just like the t* we're used to seeing, except it now deals with mean differences and the standard error of the mean differences. The overall structure of the formula remains intact. For sample sizes that are greater than 30, we would use the "inf" row (infinity) in the t-table or the normal distribution table (since the values from the "inf" row come from the normal distribution table).

According to our original claim, we're interested in **any** difference between the tires. This means that we're looking at a two-sided test. Using the t-table, we can determine the critical region boundaries with 5 (n-1) degrees of freedom and alpha (0.05) divided in two (because it's a two-sided test). Our table tells us that the region boundaries are $t_{(5,\ 0.025)} = \pm 2.57$.



Using the modified t-stat formula to find the t*, and employing the values we have already calculated, the equation yields: t* = (6.33 - 0) / (5.13 / √6) = **3.03**

Therefore, we can conclude that since t* is in the critical region, there's enough evidence to reject Ho. The alternate hypothesis is therefore supported, meaning that there is a difference between the two tires. You had better get those tires rotated, score another one for the mechanics!!

Slide 6:

Examples

Problem 1:

Construct a 95% confidence interval to estimate the difference in the paired data for tire wear given that the mean difference is 6.3 and $s_d = 5.1$.

Show Answer

Answer: Lower limit is 0.95, upper limit is 11.65.

LCL = $\bar{X} - t_{(5,0.025)} \times s/\sqrt{n} = 6.3 - 2.571(5.1/\sqrt{6}) = 6.3 - 2.571(2.0923) = 6.3 - 5.3772 =$ **0.95**
UCL = $\bar{X} + t_{(5,0.025)} \times s/\sqrt{n} = 6.3 + 2.571(5.1/\sqrt{6}) = 6.3 + 2.571(2.0923) = 6.3 + 5.3772 =$ **11.65**

This wide interval is due both to the small sample size and the large standard deviation. Recall

from the section on confidence intervals that if we were to increase "n", the interval would decrease in size, thus becoming more precise.

Slide 7:

Examples (Cont'd)

Problem 2:

The corrosive effects of various soils on coated and uncoated steel pipe were tested by using a dependent sampling plan. The data collected was summarised by:

$$n = 40,\ \sum d = 220,\ \sum d^2 = 6222$$

where d is the difference in the amount of corrosion between the coated portion and the uncoated portion of the pipe (uncoated - coated corrosion). Does this sample provide sufficient evidence to conclude that coating the pipe **makes a difference** at 99% confidence?

Show Answer

Answer: Ho: $\mu d = 0$ Ha: $\mu d \neq 0$

This is a TWO-SIDED test. Recall that when the sample size is greater than 30 for a dependent means test we use the inf. (short for infinity) row to find our critical value (or the normal distribution table).

$\alpha = 0.01/2 = 0.005$ $t_{(\text{inf},\ 0.005)} = 2.576$

$\sum d = 220$, $\sum d^2 = 6222$, $(\sum d)^2 = 48400$

$\bar{X}d = 220 / 40 =$ **5.5**

$Sd = \sqrt{[6222 - (48400 / 40)) / 39]} = \sqrt{128.5128} =$ **11.3363**

$t^* = (5.5 - 0) / (11.3363/\sqrt{40}) = 5.5 / 1.7924 =$ **3.07**

We reject Ho and accept the alternate hypothesis that there is a difference in the corrosion of treated and untreated pipes, since, t* (3.07) falls in the critical region (boundary = 2.576).

Slide 8:

Examples (Cont'd)

Problem 3:

Mark claims that his right grip strength is greater than that of his left. To test this claim, he decides to measure the grip strength of both limbs. He does a maximal contraction with each hand once a day for a week and collects the following data:

Strength (In Newtons)

Hand	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Left	46	49	46	44	50	51	47
Right	48	49	49	48	53	49	50

Set up the appropriate hypothesis test to determine if Mark's claim is correct at the 95%

confidence level. Determine the test statistic, the critical region(s), and the p-value.

Show Answer

Answer:

Since Mark wants to know if his right hand is greater than his left (not different), this is a one-sided test to the right (greater-than). First we must find the individual differences between the different trials on the different days:

Strength (In Newtons)

Hand	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Left	46	49	46	44	50	51	47
Right	48	49	49	48	53	49	50
Difference	2	0	3	4	3	-2	3

The average difference is therefore: $\sum d/n = 13 / 7 = 1.857$

Standard deviation = $\sqrt{[(51 - (13^2/7)) / 6]} = 2.1157$

- Ho: $\mu d \leq 0$, Ha: $\mu d > 0$
- critical region: $t(6, 0.05) = 1.943$
- test statistic = $(1.857 - 0) / (2.1157 / \sqrt{7}) = 2.322$
- p-value = $P(t > 2.32) = 1-0.9898 = 0.0102$
- conclusion: since t* falls in the critical region, and the p-value is smaller than 0.05, we have enough evidence to reject the null hypothesis and conclude that Mark's right grip strength is greater than his left's.

Slide 9:

Recap

If two samples are dependent, then they are typically related in some way. This usually means that although there are two sets of measurements, they are all done on the same samples. In other words, it's a "before-after" type of scenario. This means that a hypothesis test of this nature must be approached slightly differently.

- Instead of computing statistics for each sample, one column is subtracted from the other to form a column of "differences". All calculations are done using these values.
- The null hypothesis will compare the mean differences between the two samples (ud).
- Once the calculations have been done between the column values, this version of the hypothesis test is carried out the same way as the simple one.
- Since sample sizes are usually small, dependent means testing usually employs the T-distribution table. Should n > 30, use row "inf." to find the correct critical value (or the normal distribution table).

You can post a message on line in your discussion folder any time you have something to share with your discussion group concerning the current lesson. Simply click Discussion Board or use the menu at the top of the screen.

Next lesson: Independent Means Hypothesis Testing